

## CLAIMS

1. A digital data processing device, comprising:
  - a memory, said memory containing a page table, said page table having a plurality of page table entries corresponding to addressable pages, wherein each of a plurality of said page table entries contains, for each of a plurality of cacheable sub-units of the corresponding addressable page, respective cacheable sub-unit reference history data;
  - at least one processor;
  - at least one cache for temporarily storing data from said memory; and
  - a pre-fetch engine, said pre-fetch engine pre-fetching selective cacheable sub-units of addressable pages to said at least one cache, said pre-fetch engine selecting cacheable sub-units of addressable pages for pre-fetching using said cacheable sub-unit reference history data.
2. The digital data processing device of claim 1, wherein said pre-fetch engine identifies a page for selective pre-fetching upon the occurrence of at least one pre-defined event with respect to said page, and responsive to identifying a page for selective pre-fetching, selects cacheable sub-units of the identified page for pre-fetching using said cacheable sub-unit reference history data.
3. The digital data processing device of claim 2, wherein said pre-defined event with respect to said identified page comprises loading address translation data for said page into an address translation cache structure.
4. The digital data processing device of claim 3, wherein said address translation cache structure is an effective-ro-real address translation table (ERAT).

1           5.       The digital data processing device of claim 1, wherein said reference history data  
2 comprises an up-or-down counter which is incremented on the occurrence of at least one pre-  
3 defined event of a first set and decremented on the occurrence of at least one pre-defined  
4 event of a second set.

1           6.       The digital data processing device of claim 5, wherein said up-or-down counter is  
2 incremented if a respective cacheable sub-unit is loaded to said cache and referenced while  
3 in said cache, and decremented if the respective cacheable sub-unit is loaded to said cache  
4 and not referenced while in said cache.

1           7.       The digital data processing device of claim 1, wherein said digital data processing  
2 device comprises a plurality of caches at a plurality of cache levels, and wherein said pre-  
3 fetch engine pre-fetches data to different cache levels.

1           8.       The digital data processing device of claim 1, wherein said digital data processing  
2 system comprises:  
3               a plurality of processors;  
4               a plurality of caches at a plurality of cache levels, wherein a first of said cache levels  
5 comprising a plurality of caches, each cache associated with a single respective processor,  
6 and wherein a second of said cache levels comprises at least one cache which is shared by  
7 a plurality of processors.

1           9.       The digital data processing device of claim 1, wherein said digital data processing  
2 system further comprises:  
3               an address translation mechanism which translates effective addresses in an address  
4 space of a task executing on said at least one processor to virtual addresses in a global  
5 address space of said digital data processing system, and translates said virtual addresses to  
6 real addresses corresponding to physical memory locations in said memory.

1           10.     The digital data processing device of claim 9, wherein said pre-fetch engine further  
2 pre-fetches address translation data into at least one address translation cache structure of  
3 said address translation mechanism.

1           11.     A method for pre-fetching data to at least one cache a digital data processing device,  
2 comprising the steps of:

3                 maintaining, for each of a plurality of memory pages in a memory of said digital data  
4 processing device, at least one respective up-or-down counter, each said counter being  
5 incremented upon the occurrence of an event of a first type with respect to data in the  
6 memory page corresponding to the counter, and decremented upon the occurrence of an event  
7 of a second type with respect to data in the memory page corresponding to the counter; and

8                 determining whether to pre-fetch data in a page of memory based on a value of said  
9 at least one up-or-down counter corresponding to the memory page.

1           12     The method for pre-fetching data of claim 11,

2                 wherein said step of maintaining at least one up-or-down counter comprises  
3 maintaining, for each said memory page, a plurality of respective up-or-down counters in a  
4 page entry in a page table corresponding to the memory page, each up-or-down counter  
5 corresponding to a respective cacheable sub-unit of the memory page represented by the  
6 corresponding page entry in said page table, each said counter being incremented upon the  
7 occurrence of an event of said first type with respect to data in the cacheable sub-unit  
8 corresponding to the counter, and decremented upon the occurrence of an event of said  
9 second type with respect to data in the cacheable sub-unit corresponding to the counter; and

10                wherein said step of determining whether to pre-fetch data in a page of memory  
11 comprises determining whether to pre-fetch a cacheable sub-unit of the page in memory  
12 based on a value of the up-or-down counter corresponding to the cacheable sub-unit.

1 13. The method for pre-fetching data of claim 12, wherein said cacheable sub-unit  
2 corresponds to the size of a cache line of at least one cache of said digital data processing  
3 device.

1 14. The method for pre-fetching data of claim 12, wherein said step of determining  
2 whether to pre-fetch data in a page comprises:

3 identifying a page for selective pre-fetching upon the occurrence of at least one event  
4 of a third type with respect to said page;

5 responsive to identifying a page for selective pre-fetching, selecting cacheable sub-  
6 units of the identified page for pre-fetching using said cacheable sub-unit reference history  
7 data.

1 15. The method for pre-fetching data of claim 14, wherein said pre-defined event with  
2 respect to said identified page comprises loading address translation data for said page into  
3 an address translation cache structure.

1 16. The method for pre-fetching data of claim 11,  
2 wherein said event of said first type comprises loading said data into a cache of said  
3 digital data processing device, and referencing said data while it is in the cache; and.

4 wherein said event of said second type comprises loading said data into a cache of  
5 said digital data processing device, and casting said data out of said cache before it is  
6 referenced.

1 17. A processor for a digital data processing device, comprising:  
2 an instruction unit determining instruction sequences;  
3 an execution unit executing instructions in said instruction sequences;  
4 at least one cache for temporarily storing data from a memory of said digital data  
5 processing device for use by at least one of said instruction unit and said execution unit;  
6 a pre-fetch engine, said pre-fetch engine pre-fetching selective cacheable sub-units  
7 of addressable pages of said memory to said at least one cache, said pre-fetch engine  
8 selecting cacheable sub-units of addressable pages for pre-fetching using cacheable sub-unit  
9 reference history data obtained from said memory; and  
10 reference history maintenance logic, said reference history maintenance logic  
11 updating said reference history data corresponding to each respective cacheable sub-unit in  
12 response to reference history events occurring in said processor with respect to the respective  
13 cacheable sub-unit.

1 18. The processor for a digital data processing device of claim 17, wherein said pre-fetch  
2 engine identifies an addressable page for selective pre-fetching upon the occurrence of at  
3 least one pre-defined event with respect to said addressable page, and responsive to  
4 identifying an addressable page for selective pre-fetching, selects cacheable sub-units of the  
5 identified addressable page for pre-fetching using said cacheable sub-unit reference history  
6 data.

1 19. The processor for a digital data processing device of claim 18, wherein said pre-  
2 defined event with respect to said identified addressable page comprises loading address  
3 translation data for said page into an address translation cache structure.

1 20. The processor for a digital data processing device of claim 19, wherein said address  
2 translation cache structure is an effective-to-real address translation table (ERAT).

1           21.     The processor for a digital data processing device of claim 17, wherein said reference  
2           history data comprises an up-or-down counter which is incremented on the occurrence of at  
3           least one pre-defined event of a first set and decremented on the occurrence of at least one  
4           pre-defined event of a second set.

1           22.     The processor for a digital data processing device of claim 21, wherein said up-or-  
2           down counter is incremented if a respective cacheable sub-unit is loaded to said cache and  
3           referenced while in said cache, and decremented if the respective cacheable sub-unit is  
4           loaded to said cache and not referenced while in said cache.

1           23.     The processor for a digital data processing device of claim 17, wherein said pre-fetch  
2           engine further pre-fetches address translation data into at least one address translation cache  
3           structure of an address translation mechanism.

1           24.     A digital data processing device, comprising:  
2                     at least one processor;  
3                     a memory, said memory containing a page table, said page table having a plurality  
4           of page table entries corresponding to addressable pages, wherein each of a plurality of said  
5           page table entries contains persistent reference history data with respect to the corresponding  
6           addressable page, said persistent reference history data being maintained throughout the life  
7           of the corresponding addressable page in memory;  
8                     at least one cache for temporarily storing data from said memory; and  
9                     a pre-fetch engine, said pre-fetch engine pre-fetching data from said addressable  
10          pages to said at least one cache, said pre-fetch engine selecting data for pre-fetching using  
11          said persistent reference history data.

1           25.     The digital data processing device of claim 24, wherein said digital data processing  
2           system comprises a plurality of processors, and said persistent reference history data is based  
3           on reference events occurring in said plurality of processors.

1           26     The digital data processing device of claim 24,  
2                     wherein each said page table entry contains, for each of a plurality of cacheable sub-  
3                     units of the corresponding addressable page, respective cacheable sub-unit reference history  
4                     data; and

5                     wherein said pre-fetch engine determines whether to pre-fetch a cacheable sub-unit  
6                     of the page in memory based on the respective cacheable sub-unit reference history data..

1           27.     The digital data processing device of claim 24, wherein said persistent reference  
2                     history data comprises an up-or-down counter which is incremented on the occurrence of at  
3                     least one pre-defined event of a first set and decremented on the occurrence of at least one  
4                     pre-defined event of a second set.

1           28.     The digital data processing device of claim 27, wherein said up-or-down counter is  
2                     incremented if data in a respective addressable page is loaded to said cache and referenced  
3                     while in said cache, and decremented if data in the respective addressable page is loaded to  
4                     said cache and not referenced while in said cache.

1           29.     The digital data processing device of claim 24, wherein said digital data processing  
2                     device comprises a plurality of caches at a plurality of cache levels, and wherein said pre-  
3                     fetch engine pre-fetches data to different cache levels.

1           30.     The digital data processing device of claim 24, wherein said digital data processing  
2                     system further comprises:

3                     an address translation mechanism which translates effective addresses in an address  
4                     space of a task executing on said at least one processor to virtual addresses in a global  
5                     address space of said digital data processing system, and translates said virtual addresses to  
6                     real addresses corresponding to physical memory locations in said memory.

1 31. The digital data processing device of claim 30, wherein said pre-fetch engine further  
2 pre-fetches address translation data into at least one address translation cache structure of  
3 said address translation mechanism.

1 32. A digital data processing device, comprising:  
2 a plurality of processors;  
3 a memory, said memory containing a page table, said page table having a plurality  
4 of page table entries corresponding to addressable pages, wherein each of a plurality of said  
5 page table entries contains persistent reference history data with respect to the corresponding  
6 addressable page, said persistent reference history data containing reference history from said  
7 plurality of processors;  
8 at least one cache for temporarily storing data from said memory; and  
9 a pre-fetch engine, said pre-fetch engine pre-fetching data from said addressable  
10 pages to said at least one cache, said pre-fetch engine selecting data for pre-fetching using  
11 said persistent reference history data.

1 33. The digital data processing device of claim 32,  
2 wherein each said page table entry contains, for each of a plurality of cacheable sub-  
3 units of the corresponding addressable page, respective cacheable sub-unit reference history  
4 data; and  
5 wherein said pre-fetch engine determines whether to pre-fetch a cacheable sub-unit  
6 of the page in memory based on the respective cacheable sub-unit reference history data..

1 34. The digital data processing device of claim 32, wherein said persistent reference  
2 history data comprises an up-or-down counter which is incremented on the occurrence of at  
3 least one pre-defined event of a first set and decremented on the occurrence of at least one  
4 pre-defined event of a second set.